Multivariate Prediction of Total Water Storage Changes Over West Africa from Multi-Satellite Data

Ehsan Forootan, Jürgen Kusche, Ina Loth, Wolf-Dieter Schuh, Annette Eicker, Joseph Awange, Laurent Longuevergne, B. Diekkrüger, Michael Schmidt, C.K. Shum

Multivariate prediction of total water storage changes over West Africa from multi-satellite data

Ehsan Forootan forootan@geod.uni-bonn.de (1) Jürgen Kusche (1) Ina Loth (1) Wolf-Dieter Schuh (1)
Annette Eicker (1) Joseph Awange (2) Laurent Longuevergne (3) Bernd Diekkrüger (4)
Michael Schmidt (5) C. K. Shum (6) (7)

1. Institute of Geodesy and Geoinformation, Bonn University, Nussallee 17, 53115, Bonn, NRW, Germany
2. Western Australian Centre for Geodesy and The Institute for Geoscience Research, Curtin University, Perth, Australia
3. Géosciences Rennes, UMR CNRS 6118, Université de Rennes1, Rennes, France
4. Hydrology and Environmental Modelling, Department of Geography, Bonn University, Bonn, NRW, Germany
5. German Geodetic Research Institute (DGFI), Munich, Germany
6. Division of Geodetic Science, School of Earth Sciences, Ohio State University, Columbus, Ohio, USA
7. Institute of Geodesy and Geophysics, Chinese Academy of Sciences, Beijing, China

Abstract West-African countries have been exposed to changes in rainfall patterns over the last decades, including a significant negative trend. This causes adverse effects on water resources of the region, for instance, reduced freshwater availability. Assessing and predicting large-scale total water storage (TWS) variations is necessary for West Africa, due to its environmental, social, and economical impacts. Hydrological models, however, may perform poorly over West Africa due to data scarcity. This study describes a new statistical, data-driven approach for predicting West African TWS changes from (past) gravity data obtained from the Gravity Recovery and Climate Experiment (GRACE), and (concurrent) rainfall data from the Tropical Rainfall Measuring Mission (TRMM) and sea surface temperature (SST) data over the Atlantic, Pacific, and Indian Oceans. The proposed method, therefore, capitalizes on the availability of remotely sensed observations for predicting monthly TWS, a quantity which is hard to observe in the field but important for measuring regional energy balance, as well as for agricultural, and water resource management. Major teleconnections within these data sets were identified using independent component analysis (ICA) and linked via low-degree autoregressive models to build a predictive framework. After a learning phase of 72 months, our approach predicted TWS from rainfall and SST data alone that fitted to the observed GRACE-TWS better than that from a global hydrological model. Our results indicated a fit of 79% and 67% for the first year prediction of the two dominant annual and inter-annual modes of TWS variations. This fit reduces to 62% and 57% for the second year of projection. The proposed approach, therefore, represents strong potential to predict the TWS over West Africa up to two years. It also has the potential to bridge the present GRACE data gaps of one month about each 162 days as well as a - hopefully - limited gap between GRACE and the GRACE follow-on mission over West Africa. The presented method could

E. Forootan
Institute of Geodesy and Geoinformation, Bonn University, Nussallee 17, D53115, Bonn, NRW, Germany
Tel.: +49228736423
Fax: +49228733029
E-mail: forootan@geod.uni-bonn.de

also be used to generate a near-real time GRACE forecast over the regions that exhibit strong teleconnections.

Keywords Predicting GRACE-TWS · West Africa · Autoregressive model · ICA · GRACE gap filling

1 Introduction

West African climate is highly variable, ranging from tropical to semi-arid and arid over a limited 1500-km North-South gradient. The main source of precipitation over a large part of West Africa (WA) is driven by the WA monsoon system and tightly linked to large-scale pattern of ocean-atmosphere-land interaction (Giannini et al., 2003, 2008). Inter-decadal rainfall decrease over WA was highlighted as one of the largest precipitation patterns on the planet over the last half century (Ali and Lebel, 2009), leading to high risks of prolonged droughts, as in the 1970s and 1980s. Moreover, global warming adds up multiple threats to the region, with the duration and magnitude of droughts and floods expected to increase (Nicholson, 2000, Speth et al., 2011). It is of critical importance to understand and predict the impact of the WA climatic system on water resources over timescales of several months, as the livelihoods of roughly 70% of the region's population depend on uncertain rainfall and exposure to climate risk (Hansen et al., 2011).

Drought severity is classically expressed in terms of the Palmer drought index, based on moisture data only (Heim 2002). However, this index does not explicitly account for the state of all water storage compartments (Long et al., 2013), as prolonged drought conditions may have an impact on deeper groundwater systems even with limited anthropic pumping (Chen et al., 2010). Houborg et al. (2012) showed significant interest in incorporating total water storage observations (TWS), defined as the sum of all available water storage on and below the surface of the Earth, to be used for drought monitoring. Its applicability is due to the fact that TWS can represent all available forms of water resource (Scanlon et al., 2012, Schuol et al., 2008), thus, might be a better representative of drought compared to soil moisture or groundwater compartment alone.

Importance of quantifying TWS variations goes beyond its application in water resource studies. In general, the internal states of storage compartments determine their reaction to imposed boundary conditions. Runoff is driven by water stored in soil compartment and groundwater systems. Soil moisture layers - and groundwater to a lesser extend - also control evapotranspiration, cooling the land surface and regulating local energy and water balances (Koster et al., 2004). In this sense, WA has been highlighted as a 'hot spot', where the land-atmosphere coupling could play an important role, through the recycling of precipitation and the modulation of rainfall gradients (Douville et al., 2006). Main processes affecting rainfall and water availability in WA at seasonal to decadal time scales have been extensively studied within the framework of international efforts under the AMMA[1] initiative (e.g. Redelsperger et al. 2006). Beside land-atmosphere coupling, WA monsoon variability coincides with other overlapping shifts like those in global temperature and natural sea surface temperature (SST) oscillations in all tropical oceans,

[1] African Monsoon Multidisciplinary Analysis

showing remote (Pacific) or local (Atlantic and Indian) influences (see Rodríguez-Fonseca et al., 2011 and references therein, Mohino et al., 2011). Diatta and Fink (2014) studied the relationships between climate indices and monsoon rainfall, derived from rain gauge data, over West Africa, covering 1921 to 2009, and reported positive correlations between Sahel rainfall and the Atlantic Multi-decadal Oscillation (AMO), as well as the Atlantic Meridional Mode (AMM). Their results also indicated a significant impact of ENSO on inter-annual variability of precipitation over WA. Up to now, however, complex coupled ocean-atmosphere models represented limited skills to accurately simulate the main SST-WA monsoon teleconnections, both at inter-annual and decadal scales. This is reported to be caused by the simplification of different aspects of the climate system and of persistent biases (e.g., Rodríguez-Fonseca et al., 2011).

Land surface models (LSMs) and hydrological models are commonly applied to simulate the impact of climate on storage compartments (e.g., Döll et al., 2003, Rodell et al., 2004, van Dijk et al., 2013). However, the quality of the models strongly depends on model structure, boundary conditions (rainfall and evapotranspiration) and data availability, and also on model calibration/parametrization (Güntner et al., 2007). Over WA, modeling the impact of the monsoon on water resources is restricted by limited data for calibration/validation purposes (Boone et al., 2009, Schuol and Abbaspour, 2006), leading to large magnitude of uncertainties on the water balance and TWS.

Time-variable gravity solution of the Gravity And Climate Experiment (GRACE) mission offers an opportunity to remotely measure large-scale TWS changes on regional and global scale (Tapley et al. 2004, Schmidt et al., 2008a). A few studies have highlighted the critical interest of GRACE-TWS observations in WA due to the sparse distribution of in situ observation network with respect to the size of the region (Xie et al., 2012). Nahmani et al. (2012) showed that GRACE accurately estimates the annual variability of WS over WA, when compared to the output of hydrological models and GPS observations. Grippa et al. (2011) carried out a model comparison study between various GRACE products and nine land surface models (LSMs) and showed substantial differences between GRACE-TWS and LSMs. The differences were mainly ascribed to the weakness of the LSMs to correctly simulate water in surface reservoirs and evapotranspiration during the dry seasons.

This study presents a multivariate statistical TWS forecasting approach for West Africa (WA). Our goal is to capitalize on the availability of homogeneously processed, remotely sensed observations of gravity from GRACE, sea surface temperature (SST) from satellite data, as well as rainfall data from the Tropical Rainfall Measuring Mission (TRMM), and predict large-scale annual and inter-annual variability of West-African TWS changes up to a few years. Therefore, the term 'prediction' or 'forecast' in this study refers to estimation of the TWS quantity, for the period that TWS has not been observed, using its indicators, which in our case are SST and precipitation changes. A statistical approach, based on 'system identification' framework (Ljung, 1987), is chosen here for our predictions (see Section 3) since we are interested in accurate final monthly values of TWS rather than exploring the mechanism of changes in TWS compartments (e.g., soil moisture and groundwater) and their interactions. A similar concept has already been used, e.g., by the USA's National Oceanic and Atmospheric Administration (NOAA) for predicting climatic parameters (http://www.cpc.ncep.noaa.gov/). We should

mention here that one could also alternatively use 'gap-filler' approaches (e.g., Rietbroek et al., 2014) to estimate (or predict) surface load or TWS over a region of interest. We will show later that the prediction approach here provides up to two years TWS predictions, while retaining the spatial resolution of GRACE products. Reager and Famiglietti (2013) presented an experimental predictions approach that relates water storage changes to precipitation forcing and then generalize the relation based on large-scale basin characteristics. Unlike our proposed method, this approach requires extra information about basin characteristics.

The predictability skill of the proposed statistical approach would be optimal if major physical processes over the region of study are included in the learning phase. Thus, both ocean-atmosphere and land-atmosphere processes are represented as predictors of West African TWS changes: (i) SST variations over the major oceanic basins of the Atlantic, Pacific, and Indian Oceans (ii) TRMM rainfall observation over West Africa. Although other predictors might also improve the forecasting results (evapotranspiration, soil moisture changes), model quality is related to parsimony and data homogeneity. Here, we only rely on SST and rainfall data since they are more accurately derived from remote sensing observations compared to the other possible indicators (see, e.g., Reynolds et al., 2002, Huffman and Bolvin, 2012, Wang and Dickinson, 2012).

Time-variable maps of predictors (SST and rainfall data) and predictands (TWS data) include large temporal and spatial correlations. This requires application of a dimension reduction method before constructing the mathematical relationship between predictors and predictands (e.g., Kaplan et al., 1997). This considerably improves the skill of the forecasting approach (see e.g., Westra et al., 2008). The statistical method of independent component analysis (ICA) was applied to extract individual modes of variability that are mutually independent and successively explain the maximum amount of existing variance in the data (Forootan and Kusche, 2012, 2013). An optimum autoregressive model with exogenous variables (ARX) (Ljung, 1987) was then used to relate independent components (ICs) of predictands to ICs of predictors. In the end, the model allows a thorough representation of complex processes in a highly efficient way as compared to physical models. The combination of ICA/ARX modeling can be generalized worldwide, with an adequate identification of likely forcing recalibration of the model, with respect to the region of interest. Examples include the regions such as North America and the Australian continent, which exhibit strong ocean-land-atmosphere interactions (Douville et al., 2006, Forootan et al., 2012).

From a methodological point of view, we prefer the ICA algorithm for dimension reduction over, e.g., principal component analysis (PCA, Preisendorfer, 1988); this view is rooted in the improved performance of ICA in extracting trends, annual patterns, as well as slow dynamic patterns such as the El Niño-Southern Oscillation (ENSO) from climate observations (e.g., Aires et al., 2002, Ilin et al., 2005)). In a preliminary study, Forootan and Kusche (2012) and Forootan et al. (2012) applied the ICA method to global and local GRACE-TWS time series, and demonstrated its value in extracting climate related patterns. We also evaluated the use of PCA in our proposed statistical TWS forecast (results are not shown here), and found that ICA improves the extraction of teleconnections, e.g., ENSO and the Indian Ocean Dipole (IOD) patterns, as well as the performance of the prediction. A similar conclusion was reached, e.g., by Westra et al. (2007), who assessed the performance of ICA and PCA for simulating hydrological time series.

Finally, we prefer the ARX model over the common canonical correlation analysis (CCA) approach (e.g., von Storch and Navara, 1999) for relating predictors and predictands since ARX offers more flexibility to relate multiple parameters as exemplified e.g., in Westra et al. (2008).

To implement our prediction approach, first, we begin by decomposing the following data sets individually into statistically independent modes: (i) GRACE-TWS changes over West Africa, to provide the dominant independent patterns of total water storage (TWS) that are subsequently identified with the predictands of the ARX process; (ii) SSTs over the Atlantic, Pacific, and Indian Ocean basins, in order to extract ocean-atmospheric interactions and teleconnections; and (iii) TRMM data over West Africa, for extracting the main patterns of rainfall over the region. Then, the modes found by analyzing (ii) and (iii) are introduced as predictors of (i) within the ARX process model, while using the first 72 months of (i), (ii) and (iii) for the training step (statistical simulation). The fitted ARX model, along with the independent modes of SST and rainfall (i.e. the predictors of the ARX model) after the 72'th month are then used to predict TWS after the simulation period. The prediction is evaluated using those GRACE-derived TWS anomalies that are available after the simulation period. Forecasting error levels are also predicted using a Monte Carlo error estimation process. We should mention here that to decompose water storage and rainfall data in (i) and (iii), we introduced the West Africa region as a simple box (latitude between 0° to $25^\circ N$ and longitude between -20° to $10^\circ E$). The method, however, can be extended to grids delineated by basin shape or basin-averaged time series.

This contribution is organized as follows; in Section 2, we briefly present the data sets used in the study. The dimension reduction and the ARX forecasting methods are introduced in Section 3, followed in Section 4 by a discussion of the leading independent modes found in the total water storage (TWS), sea surface temperature (SST) and rainfall data sets. In Section 5, we discuss the results of ARX-TWS simulations and forecasts over West Africa. The study is concluded in Section 6. The paper also contains two appendices. In Appendix A, the details of GRACE-TWS estimations over West Africa are described, and in Appendix B, we present the details of mathematical methods, used in this study, including ICA, and the ARX model as well as their uncertainty estimations.

2 Data

2.1 Total Water Storage from GRACE and WGHM

The GRACE mission consists of two low-earth orbiting satellites in the same orbital plane at the current altitude of ~ 450 km and in an inclination of 89.5°. The separation distance between the two satellites is measured precisely by a K-band ranging system and the location of each satellite is determined by GPS receivers on-board the spacecraft (Tapley et al., 2004). These data, after application of several corrections, are then used by a number of analysis to generate time-variable (usually monthly) Level-2 gravity field products (Flechtner, 2007). In this study, we used monthly GRACE products from the German Research Centre (GFZ) Potsdam (Flechtner, 2007) for computing TWS fields, covering August 2002 to May 2011. We did not interpolate the missing data of January 2003, 2004, May

2003, and December 2008 in order to avoid creating artifacts. For comparison, we also used monthly GRACE-ITG2010 products from Bonn University, Germany (Mayer-Gürr et al., 2010), which are available for the period between September 2002 and August 2009, and are provided together with full variance-covariance information. The covariance matrices were used to estimate the accuracy of the GRACE-TWS grids. Total water storage from GRACE are also compared with TWS output from the WGHM model (Döll et al., 2003), covering the years 2003 to 2010. WGHM represents the major hydrological components, such as soil moisture, rainfall, snow accumulation, melting, evaporation, runoff, and the lateral transport of water within river networks. For this study, we prefer WGHM over using land surface models (LSM) since it also contains a groundwater simulation model and, therefore, its vertically aggregated storage can be directly compared to GRACE TWS. The details of data preparation are described in Appendix A.

2.2 SST

Monthly reconstructed global $1^\circ \times 1^\circ$ Reynolds sea surface temperature (SST) data (Reynolds et al., 2002) were used over the period 2002 to 2012. The Reynolds SST has been frequently used for climate studies, including some addressing African rainfall variability in relation to SST (e.g., Mohino et al., 2011, Omondi et al., 2012). Similar to Omondi et al. (2012), the SST data cover three major ocean basins: we include an Atlantic Ocean box (-66° to 13° E and -20° to 31° N), a Pacific Ocean box (159° to 275° E and -30° to 19° N) and an Indian Ocean box (34° to 114° E and -50° to 1° N). Sea surface temperatures in these regions are then extracted and analyzed through the ICA approach (Section 4). We found that a slight difference in the size of the selected boxes would not change the results of ICA significantly.

2.3 TRMM

Version 7 of TRMM-3B42 products (Huffman and Bolvin, 2012) covering 2002 to 2012 (http://mirador.gsfc.nasa.gov/) was used. The downloaded 3-hourly rainfall rates have been converted to rainfall amount and aggregated to monthly basis. TRMM was previously used e.g., by Nicholson et al. (2003) to study the patterns of precipitation over West Africa. Huffman and Bolvin (2012) and Fleming and Awange (2013) reported a significant improvement of version 7 over version 6, likely as a result of including more microwave sounding and imagery records as well as implementing better processing algorithms.

3 Methodology

3.1 Identifying Dominant Independent Patterns from Available Data

In order to keep the problem of identifying independent modes on a grid, including the statistical relationships between them, numerically manageable, it is mandatory to first apply a dimension reduction method before constructing the

mathematical relationship between predictors and predictands (e.g., Kaplan et al., 1997). This improves the prediction skills of the statistical approach, since the redundant information within the data sets, both predictors and predictands, will be reduced. Dimension reduction is implemented here by applying a 2-step independent component analysis (ICA) algorithm (Forootan and Kusche, 2012, 2013) to the TWS, rainfall, and SST data sets. Forootan and Kusche (2012) formulate two alternative ways of applying ICA, in which either temporally independent components or spatially independent components are constructed. The temporal ICA method, which we simply abbreviate here as ICA, is preferred in this study since it provides temporally independent components that allow the development of a prediction model that is univariate in the predictand (see Section 3.2).

Let us assume that GRACE-TWS anomalies (in mm), after removing the temporal mean, are stored in the matrix $\mathbf{X}_{TWS} = \mathbf{X}_{TWS}(s,t)$, where t is the time, and s stands for spatial coordinate (grid points). Applying ICA means that $\mathbf{X}_{TWS}$ is decomposed into spatial and temporal components as

$$\mathbf{X}_{TWS} = \mathbf{Y}_j \mathbf{A}_j^T, \tag{1}$$

where the columns of $\mathbf{Y}_j$ contain the j dominant unit-less temporally independent components (ICs) of TWS, and the columns of $\mathbf{A}_j$ represent the corresponding spatial maps. Each temporal pattern (i.e. a column of $\mathbf{Y}_j$) along with the corresponding spatial pattern (a column of $\mathbf{A}_j$) represent an independent mode of variability. Likewise, the temporally centered maps of rainfall over West Africa $\mathbf{X}_{Rainfall}$, and of SST over the major oceans $\mathbf{X}_{SST}$ can separately be written as

$$\mathbf{X}_{SSTorRainfall} = \mathbf{U}_{j'} \mathbf{B}_{j'}^T, \tag{2}$$

where $\mathbf{U}_{j'}$ stores the j' dominant unit-less temporally independent components of SST or rainfall, and the columns of $\mathbf{B}'_j$ contain the associated spatial maps. We used different indices j and j' in Eqs. (1) and (2) to emphasize that the number of retained modes from different data sets are not necessarily the same. Selecting a proper subset (j or j') is addressed in Appendix B.

In our analysis, we found the spatial patterns associated with independent modes of total water storage anomalies, i.e. the columns of $\mathbf{A}_j$ (Eq. (1)) to be sufficiently stable. This means that, for instance, the spatial patterns ($\mathbf{A}_j$) derived from 10 years of TWS data do not differ significantly from those derived from 8 or 12 years of data. Therefore, for building the forecasting model, we only link the temporal components (ICs) of the predictor data sets (all columns of $\mathbf{U}_{j'}$ derived from SSTs and TRMM-rainfall) to individual ICs of the predictand (each column of $\mathbf{Y}_j$ in Eq. (1)). Finally, we will use the $\mathbf{A}_j$ derived from TWS in order to reconstruct the forecasting maps. Details of the ICA decomposition and the corresponding error estimation are addressed in Appendix B.

3.2 Prediction using an Autoregressive Model with Exogenous Variables (ARX)

An ARX process is governed by a system of linear equations, which describe the relationship between the current and previous values of the system output and the values of inputs. In our case, the ARX model is formulated as a multiple-input

(the independent modes or ICs of SST and rainfall all together) and single-output (each IC of TWS) model (Ljung, 1987):

$$y(t) + \sum_{i=1}^{n_a} a_i y(t-i) = \sum_{q=1}^{m} \sum_{l=1}^{n_b} b_{q,l} u_q(t - k_q - (l-1)) + \xi(t), \tag{3}$$

where y represents a particular mode of TWS, i.e. $y(t)$, $t = 1, \ldots, n$, represent a column of $\mathbf{Y}_j$ in Eq. (1). In Eq. (3), n_a is the order of the ARX model with respect to the predictand, $u_q(t)$, $q = 1, \cdots, m$, and $t = 1, \ldots, n$, are ICs of SSTs and rainfall from $\mathbf{U}_{j'}$ in Eq. (2), while m is the number of predictors. The order of the ARX model with respect to the predictors is n_b, and k_q denotes the number of time-steps before the q'th input (predictor) affects the output y, i.e. the dead time of the system. Finally, ξ allows for a white-noise random input. The coefficients of the ARX models a_i, $i = 1, \cdots, n_a$, and $b_{q,l}$, $q = 1, \cdots, m$, and $l = 1, \cdots, n_b$, have to be derived in the simulation step, using both predictand and predictors (Ljung, 1987). Once the coefficients are computed, in the forecasting step only the predictors (ICs of SST and rainfall) are used to estimate the values of TWS after the simulation period. Details of the computations and error propagation are addressed in Appendix B.

4 Dominant Independent Modes of TWS, SST, and Rainfall

When following the decomposition procedure as described in Section 3.1, we identify two independent, statistically significant, modes in GRACE GFZ-TWS (abbreviated as GFZ-TWS), four independent modes in SST changes over the Atlantic Ocean, three modes over the Pacific Ocean, and four modes over the Indian Ocean. For the rainfall changes, also four significant independent modes were obtained. Our approach for separating significant modes from the insignificant ones is also presented in Appendix B. Finally, Table 1 summarizes the variance percentage of the modes discussed so far.

Table 1 List of the variance percentage that each of the independent mode in Section 4 represents. Independent modes of GRACE GFZ-TWS are shown in Fig. 1, a and b, those of the Atlantic, Pacific and Indian Ocean-SST are shown in Fig. 2. Fig. 3 contains the independent modes of TRMM-rainfall.

	First Independent Mode	Second Independent Mode	Third Independent Mode	Fourth Independent Mode
GRACE GFZ-TWS	62.4%	20.4%		
WGHM-TWS	60.4%	16.4%		
Atlantic Ocean-SST	46.4%	25.4%	11.7%	4.1%
Pacific Ocean-SST	55.6%	15.4%	14.6%	
Indian Ocean-SST	47.8%	15.4%	14.3%	11.1%
TRMM-rainfall	44.2%	12.3%	6.9%	6.4%

Dominant Independent Modes Identified in Total Water Storage Variability

The first and second independent modes the from TWS anomalies derived from GRACE (GFZ solutions) are shown in Fig. 1,a. The first dominant independent mode, which explains 62.4% of variance, represents the annual water variability over West Africa. Here a damping of the signal magnitude can be seen in the year

2005 (temporal IC1 of GRACE GFZ-TWS). From the spatial pattern of IC1, a concentration of annual variability appears to be dominant over the tropic and coastal regions. The second mode of GRACE-TWS contains inter-annual variations, along with periodic components of ~ 3 and 5 years period. Nicholson (2000) found a similar period in rainfall variations over West Africa. The second independent mode represents 20.4% of the GRACE-TWS variance, thus, the first two leading modes in Fig. 1 represent more than 80% of the TWS variance over West Africs. In fact, the dominanting annual and inter-annual variability of TWS are found well separated, and thus can be treated separately within the ARX simulation and prediction steps. We attribute this behavior to the properties of the ICA decomposition method.

The separate treatments of the two annual and inter-annual variability of water storage components seems to be reasonable since they are likely due to different physical influences from environment (here the indicators SST and rainfall). Therefore, the mathematical relationship between TWS and its indicators must be separately weighted (i.e. coefficients a_i, $i = 1, \cdots, n_a$, and $b_{q,l}$, $q = 1, \cdots, m$ of IC1-GRACE and IC2-GRACE will be separately computed in Eq. 3). We would like to mention here that the oscillations that exist in the extracted dominant independent modes are not necessarily explained by a fundamental annual or inter-annual cycle and its overtones. Therefore, for the decomposition and prediction procedures, we chose not to reduce any such pre-defined oscillations (see also Schmidt et al. 2008b).

For comparison, we then projected WGHM-TWS and ITG2010-TWS on to the spatial patterns of Fig. 1,a, using Eq. (B2) (Appendix B). The results are shown by the black and gray lines in Fig. 1,a. The temporal patterns indicate that the annual TWS changes from WGHM (IC1 in Fig. 1,a) are comparable to those seen by GRACE, whereas at the inter-annual time scale they are different (see IC2 in Fig. 1,a). To confirm this finding, we also applied ICA (Eq. (1)) to the WGHM-TWS data, individually, with the results shown in Fig. 1,b. The first independent mode of WGHM-TWS (60.4% of variance) is comparable to that of GRACE-TWS, while the second mode of WGHM-TWS (spatial and temporal pattern of IC2-WGHM) that accounts for 16.4% of variance is found quite different from those of GRACE-TWS (both GFZ and ITG2010). This finding shows that WGHM-TWS and GRACE-TWS are not consistent at inter-annual time scale; compare Fig. 1,a with Fig. 1,b. We hypothesize that the difference could be attributed to the possible miss-modeling of surface water storage or water withdrawals over the region. Further research will be needed to address the exact cause of differences, but this is outside the scope of the current study.

Dominant Independent Modes of Sea Surface Temperature and Rainfall Data

Applying the ICA approach (Eq. (2)) to SST changes over the three ocean boxes shows that their first two independent modes are related to the annual variability of SST (IC1 and IC2 in Fig. 2,a,b, and c). Over the Atlantic, for instance, IC1 and IC2 are related to the annual dipole structure, which also correlate very much with IC1 of GFZ-TWS. The same damping of the annual amplitude in the year 2005 is seen for IC2-Atlantic SST, similar as with IC1 of GFZ-TWS. Our result confirms that the recent annual variability of total water storage over West Africa is highly correlated with the Atlantic ocean-atmospheric interactions, reflected in

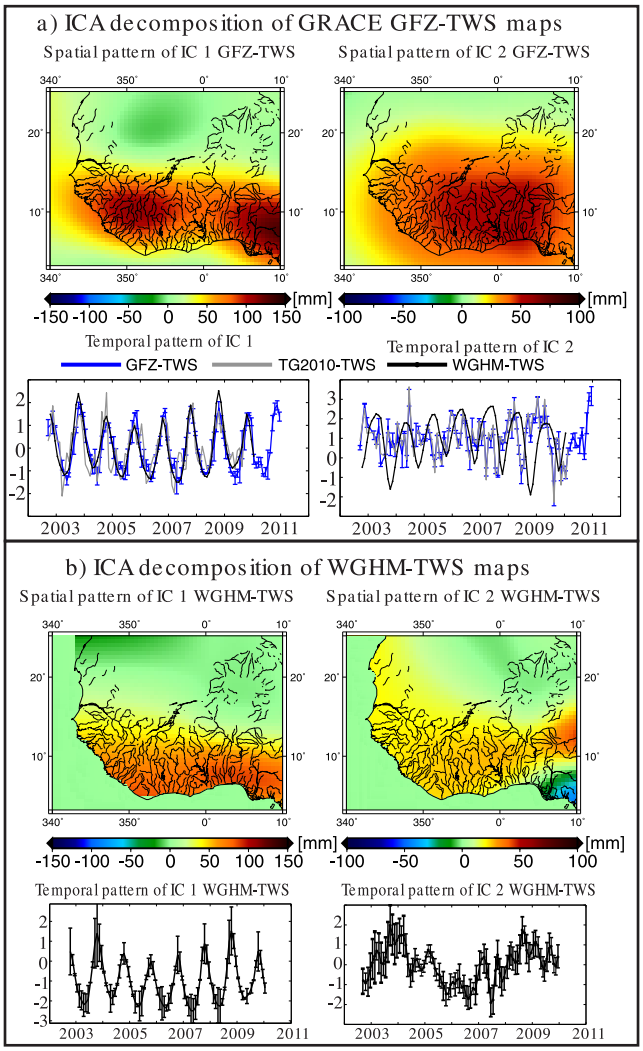

Fig. 1 Overview of the ICA decomposition of TWS changes over West Africa (counted as predictands). a) ICA decomposition of GRACE GFZ-TWS data (GFZ-TWS). For comparisons, WGHM-TWS and ITG2010-TWS changes are projected on the spatial patterns of IC1 and IC2, using Eq. B2 in Appendix B. The results are presented along with temporal ICs of GFZ-TWS. b) ICA decomposition of WGHM-TWS maps. The variance fraction of each independent mode is presented in Table 1. Uncertainties are shown by error-bars around temporal components. Details of uncertainty computations can be found in Appendix B.

the SST data (see similar findings in e.g., Mohino et al., 2011). For the variance percentages that each mode represents, we refer to Table 1.

We find that the third mode of SST changes over the Atlantic and Indian Ocean boxes represents semi-annual variability, while IC3-Pacific SST represents the ENSO pattern; we compared IC3-Pacific SST with the monthly ENSO pattern (shown by the Southern Oscillation Index (SOI)) provided by the Australian

Bureau of Meteorology (http://www.bom.gov.au/climate/enso/). A high correlation of 0.84 was obtained, suggesting that the pattern is physically meaningful (IC3 in Fig. 2,b). We also found a significant correlation of 0.68 between IC2 of GRACE-TWS (Fig. 1,a) and SOI, revealing a relationship between total water storage variability over West Africa and ENSO.

IC4-Atlantic SST represents a complicated pattern, which we do not attempt to interpret here. IC4-Indian SST (Fig. 2,c) follows the Indian Ocean Dipole (IOD) pattern (see e.g., Saji et al., 1999). Comparing our results to the IOD index derived from the Japan Agency for Marine-Earth Science and Technology (JAMSTEC, http://www.jamstec.go.jp/frcgc/research/

d1/iod/HTML/Dipole%20Mode%20Index.html) represents a significant correlation of 0.73. This shows that ICA extracts teleconnection patterns from SST data fairly well.

Four independent modes were extracted from TRMM-rainfall data, from which IC1-TRMM and IC2-TRMM relate to the annual rainfall variability with three months phase differences (Fig. 3). In 2005, a damping of the signal magnitude can be seen in IC2-TRMM, but it is less pronounced compared to that of IC2-Atlantic SST. This might be due to the fact that IC2-TRMM represents a local impact, compared to the large-scale interaction that IC2-Atlantic SST represents. IC3-TRMM and IC4-TRMM apparently represent the semi-annual rainfall variations. We found a lag of two months between the ICs of rainfall and those of TWS.

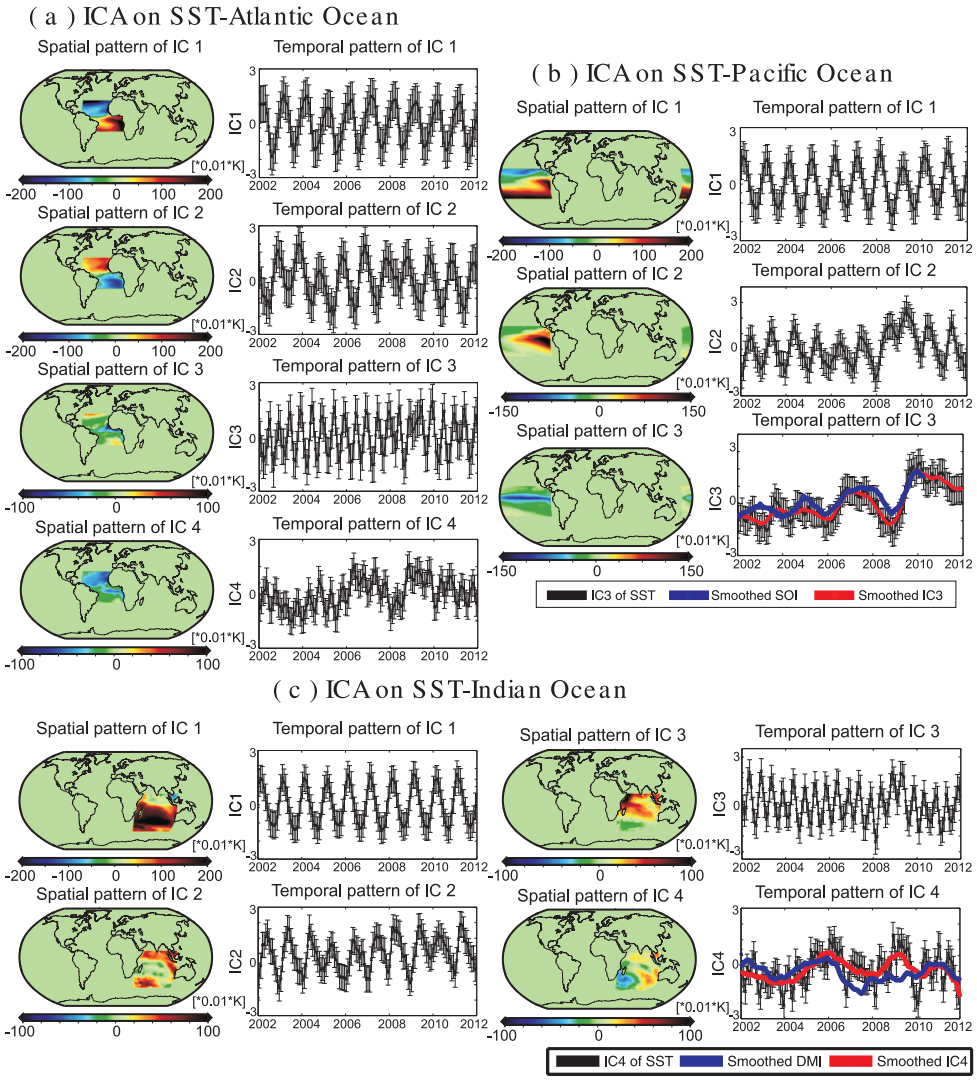

Fig. 2 Overview of ICA decomposition applied to the SST changes over (a) the Atlantic, (b) Pacific, and (c) Indian Oceans. The variance fraction of each independent mode is presented in Table 1. Uncertainties are shown by error-bars around temporal components. The smoothed temporal patterns in (a) and (b) are derived by applying a 12-month moving average filter.

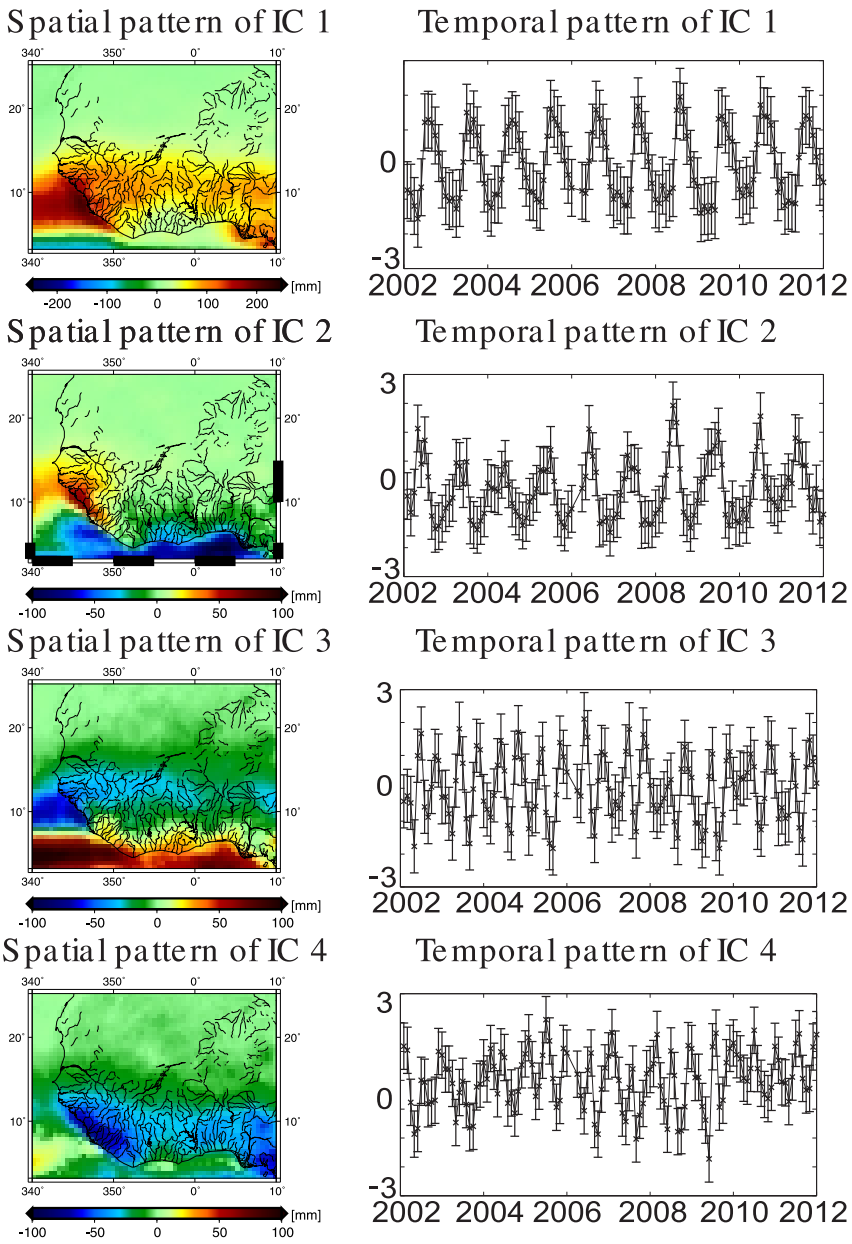

Fig. 3 Overview of ICA decomposition applied on rainfall changes over West Africa. The variance fraction of each independent mode is presented in Table 1. Uncertainties are shown by error-bars around temporal components.

5 Predicting Total Water Storage with an Autoregressive Model

Training Step

To compute the best-fitting ARX model, we have inserted the first 72 months of each mode of GRACE GFZ-TWS (IC1 and IC2 in Fig. 1,a) and the first 72 months of all temporal modes of SST and rainfall (ICs of Figs. 2 and 3) in Eq. (3). Before performing the training step of the ARX model, the data for January 2003, 2004, May 2003 and December 2008 were excluded from the input time series (ICs of SSTs and rainfall) to synchronize them with the GFZ-TWS time series. Then, we had to choose optimum n_a and n_b; these were found experimentally by varying them between one to three. The time delay k_q was searched for between zero and three months. This was then followed by running the ARX simulation step (Eq. (3)). We did not consider higher orders for n_a and n_b since we would like to keep the forecasting model as simple as possible (e.g., Westra et al., 2008). For k_q, previous studies (e.g., Ahmed et al., 2011) found a delay of up to three months between SST-rainfall and TWS changes. The required coefficients for each ARX model $\hat{\mathbf{\Theta}}$ were computed using Eq. (B5) in Appendix B.

Our numerically simulated results for both IC1 and IC2 of GFZ-TWS suggest that an ARX model with n_a=1 and n_b=3 provides the best fit with the residuals passing the normality test. The RMS of differences between the simulated TWS values from the ARX process and the ICs of GFZ-TWS (Eq. (B6) in Appendix B) was used as the fit criterion. Two sets of k_q corresponding to the simulations of the IC1 and IC2 of GFZ-TWS were found and are presented in Table 2. Our simulation results indicate that the ARX models provide a fit of 93% and 83% for simulating the two dominant components of GFZ-TWS (IC1 and IC2 in Fig. 1,a), respectively. The simulation fit of ARX corresponding to IC2 of GFZ-TWS is, however, lower than that of IC1 since its temporal pattern appears much more complicated than the annual pattern in IC1. Therefore, it might not have been fully captured by the predictors (see Fig. 4).

Table 2 Time delays k_q derived from simulation of IC1-GRACE GFZ-TWS and IC2-GRACE GFZ-TWS. The values are in month and denote the number of time-steps before each predictor (ICs of SSTs and ICs of TRMM-rainfall respectively in Figs. 2 and 3) affect the output (each individual IC of GRACE GFZ-TWS in Fig. 1, a).

	IC1 SST Atlantic Ocean	IC2 SST Atlantic Ocean	IC3 SST Atlantic Ocean	IC4 SST Atlantic Ocean	IC1 SST Pacific Ocean	IC2 SST Pacific Ocean	IC3 SST Pacific Ocean	IC1 SST Indian Ocean	IC2 SST Indian Ocean	IC3 SST Indian Ocean	IC4 SST Indian Ocean	IC1 TRMM West Africa	IC2 TRMM West Africa	IC3 TRMM West Africa	IC4 TRMM West Africa
k_q related to IC1 of GFZ-TWS	1	0	2	1	0	3	1	1	0	1	0	0	3	3	1
k_q related to IC2 of GFZ-TWS	1	0	1	1	0	0	1	1	0	1	0	1	3	1	3

To assess the sensitivity of the ARX models (Eq. (3)) with respect to each input, first, for each IC of GFZ-TWS, each IC of Figs. 2 and 3 was individually inserted in Eq. (3) and the ARX model evaluated. For each IC of GFZ-TWS, therefore, an ensemble of 15 ARX-modeled TWS outputs was generated, and the correlations of these outputs and the ICs of GFZ-TWS were then computed. From the 15 ARX-generated TWS, those that represented the largest correlations with their corresponding IC of GFZ-TWS were likely to have the most influence on the prediction. Our results show that the ARX-outputs generated by IC1, IC2-Atlantic

SST, and IC1-rainfall had the largest influences on the ARX model of IC1 GFZ-TWS. The prediction of IC2 GFZ-TWS was found to be sensitive to IC3-Pacific SST, IC3-Atlantic SST, and IC3-rainfall. The most sensitive indicators and the correlations of the associated outputs with the ICs of GFZ-TWS are presented in Table 3. One might use these results for model reduction of the original ARX process (see e.g., Westra et al., 2008), however, such a reduction was not applied in this study since (i) the fit derived from each of the 15 model run was smaller than that of original ARX run, and (ii) the ARX models apparently possessed sufficient degree of freedom to be computed based on the current indicators.

Table 3 List of the most sensitive indicators derived from the ARX models for each input and the computed correlation of the output with IC1 and IC2 of GRACE GFZ-TWS. Correlations are derived at 95% level of confidence.

Rank:	1	2	3
Model run by:	IC1 SST Atlantic Ocean	IC2 SST Atlantic Ocean	IC1 TRMM-rainfall West Africa
Correlation with IC1 of GFZ-TWS	0.81	0.61	0.52
Model run by:	IC3 SST Pacific Ocean (ENSO)	IC3 SST Atlantic Ocean	IC3 TRMM-rainfall West Africa
Correlation with IC2 of GFZ-TWS	0.59	0.46	0.42

ARX Forecasting Step and Validation

Having simulated the ARX model parameters for the two dominant independent modes of GFZ-TWS ($\hat{\boldsymbol{\Theta}}$ is known from Eq. (B5), within the training step), we used the indicator time series, i.e. all ICs shown in Figs. 2 and 3 after the 72-months training period alone to predict ICs of GRACE GFZ-TWS changes for the years 2010 and 2011. The predictions were derived from Eq. (B8), and their uncertainties were evaluated using the Monte Carlo approach described in Appendix B. Training and forecast results are shown in Fig. 4a,b. Our prediction approach, therefore, uses only the process structure described in Section 3.2 and the determined lag relation between the predictors and ICs of GFZ-TWS changes (i.e. Table 2). The fit of the forecast for the first leading mode of total water storage, when compared to the observed GFZ-TWS values, after one year was found 79%, while after two years this reduced to 62%. As Fig. 4a also shows, after two years, the standard deviation of the propagated uncertainty is quite large. This suggests that the proposed approach is more or less reliable for predictions of up to two years. Fig. 4b shows that the fit of the forecast for the second leading independent pattern of total water storage after one year was reduced to 67%. After two years, a fit of 57% was found. Comparing projected values of WGHM-TWS (black lines in Fig. 1a) with the ICs of GFZ-TWS, during the first year of forecast, we found a fit of 78%

Table 4 List of the computed fit values (Eq. B6) derived from comparing the ARX-TWS outputs with the IC1 and IC2 of GRACE GFZ-TWS. We also compared the projected values of WGHM-TWS in Fig. 1 with the IC1 and IC2 of GRACE GFZ-TWS. The values indicate that the ARX outputs are closer to that of GRACE GFZ-TWS.

	Simulation period (72 months)	First year of the forecast	Second year of the forecast
Fit values of the ARX model with respect to IC1 of GFZ-TWS:	93%	79%	62%
Fit values of the ARX model with respect to IC2 of GFZ-TWS:	83%	67%	57%
Fit values of the WGHM model with respect to IC1 of GFZ-TWS:	91%	78%	43%
Fit values of the WGHM model with respect to IC2 of GFZ-TWS:	54%	53%	31%

for the annual and a moderate fit of 53% for the inter-annual pattern. This result indicates that the TWS prediction from the proposed statistical method is indeed closer to the observed GFZ-TWS changes over West Africa, when compared to hydrological modeling. See the fit values in Table 4.

Comparing the ARX-derived TWS predictions with the ICs of GFZ-TWS, during the forecast period, we found no apparent deviations between TWS time series (see Fig. 4a,b), thus, the reported fit values are significant. For the inter-annual time-scale, however, specific care should be taken since the simulation and prediction of the ARX-TWS method is very much sensitive to the temporal patterns of the input parameters. When the ICs of SSTs or rainfall are not well defined, the inter-annual forecast of ARX-TWS might perform poorly or be biased. This has been tested by replacing the ICs of SSTs and TRMM-rainfall with temporal components derived e.g., from the principal component analysis (PCA) decomposition for running the ARX predictions (results are not shown). A bias was found in the prediction of IC2 GFZ-TWS, which was most likely due to the fact that the PCA-derived indicators (PCs of SST and rainfall) do not reflect the inter-annual TWS changes sufficiently.

In order to assess the robustness of the performed forecast with respect to the training period, we performed a backward simulation and forecast, i.e. the last 72 months of both indicators (ICs of SSTs and TRMM-rainfall) and predictands (individual ICs of GFZ-TWS) were used in the simulation step to predict the first two years of GFZ-TWS. The results, summarized in Fig. 5, show a fit similar to the forward forecast in Table 4. Therefore, even though the training step was quite short, no temporally variable bias was found in both forward and backward predictions. This confirms the robust performance of the proposed ICA-ARX approach, at least, with respect to the performed tests.

Inserting the time series of the prediction and the spatial components of Fig. 1,a in Eq. (1), one may reconstruct the TWS maps for the period when the ARX model and their inputs are available. In this case, our prediction values provide $\mathbf{Y}$ and the spatial maps of Fig 1,a provide $\mathbf{A}$ in Eq. (1). Fig. 6 compares the original GRACE GFZ-TWS values of the year 2010, after removing the temporal mean of 2003 to 2011 and the contribution of Lake Volta (as discussed in Section 3 and Appendix A), with the values of the ARX-TWS forecast. The predictions fit quite well to GFZ-TWS fields. Fig. 6,c shows the difference between GFZ-TWS and the predicted ARX-TWS. The patterns of the differences are similar to the striping

noise of GRACE solutions (Kusche, 2007). Our results, therefore, support the idea of using the presented statistical approach to forecast TWS changes over West Africa.

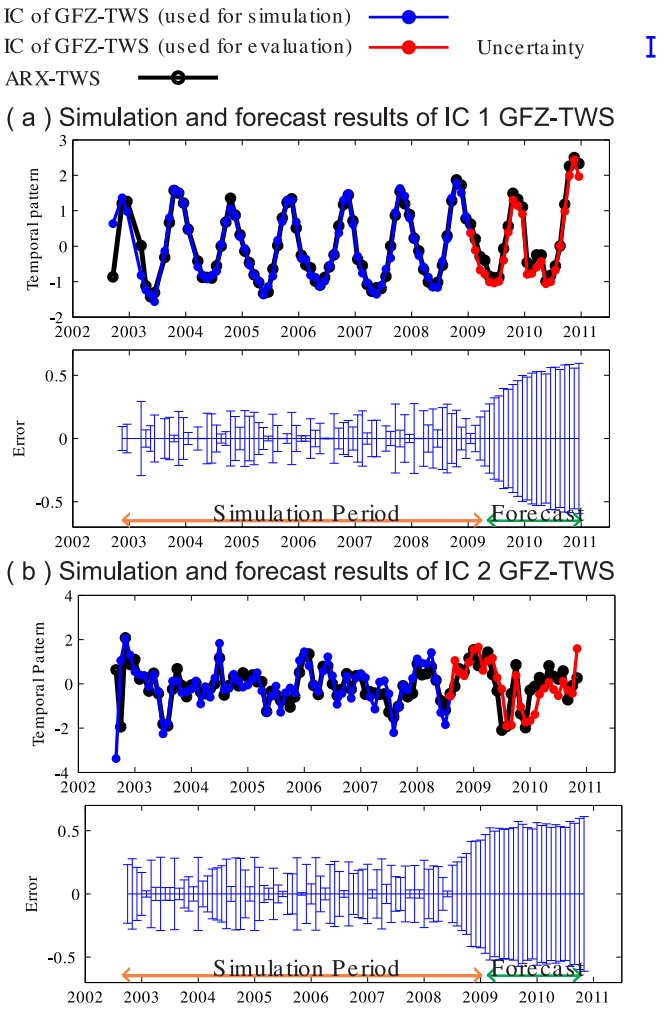

Fig. 4 Results of simulations and forecasts of IC1 and IC2 of GRACE GFZ-TWS (Fig. 1,a), using the ARX models (shown by black-lines). Fig. 4a(top) represents the results for IC1 of GRACE GFZ-TWS while assuming the ICs of Figs. 2 and 3 as indicators. Fig. 4a(bottom) shows the uncertainty of the forecast on top. Fig. 4b(top) represents the same results as (a) but corresponding to IC2 of GRACE GFZ-TWS. Fig. 4b(bottom) indicates the uncertainty of the forecast on top. For simulation, the first 72 months of TWS are used (shown in blue). TWS values after the 72th month are then used for evaluating the forecasts (shown in red). Reconstructing the forecast for the year 2010, and its comparison with the original GRACE GFZ-TWS fields are presented in Fig. 6.

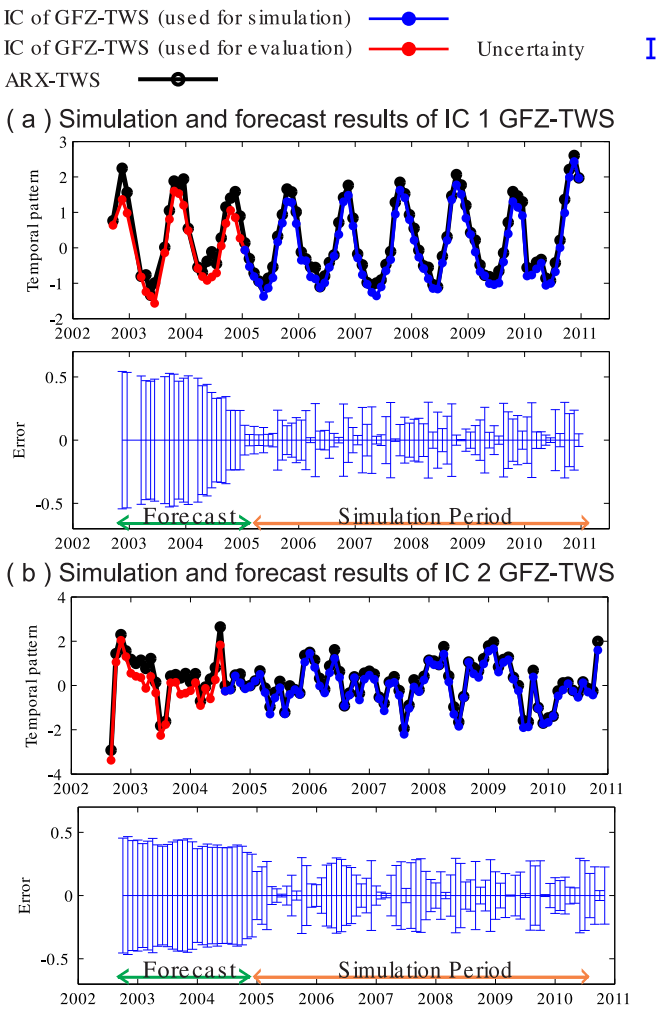

Fig. 5 Backward simulations and forecasts of IC1 and IC2 of GRACE GFZ-TWS (Fig. 1,a), using the ARX models (shown by black-lines). The results are similar to those of Fig. 4, however here, the last 72 months of both indicators (ICs of SSTs and TRMM-rainfall) and predictands (individual ICs of GRACE GFZ-TWS) are used in the simulations and the first two years of IC1 and IC2 of GRACE GFZ-TWS predicted. Fig. 5a(top) represents the results for IC1 of GRACE GFZ-TWS. Fig. 5a(bottom) shows uncertainty of the forecast on top. Fig. 5b(top) represents the same results as (a) but corresponding to IC2 of GRACE GFZ-TWS. Fig. 5b(bottom) indicates the uncertainty of the forecast on top.

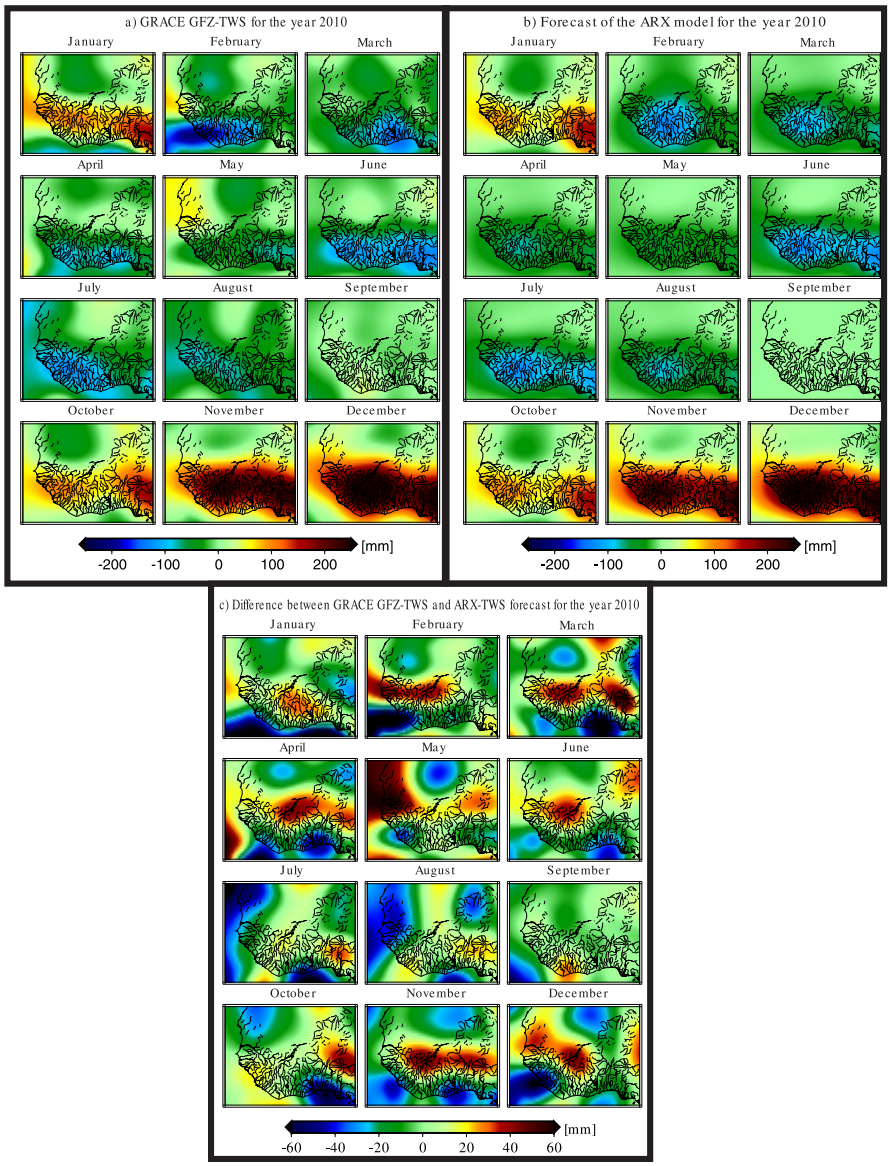

Fig. 6 Overview of TWS maps for the year 2010 over West Africa, without the signal of Volta, and after removing the temporal mean of 2003 to 2011. (a) TWS maps derived from GRACE (GFZ RL04 data), (b) TWS maps derived from the statistical forecast (the ARX models and ICA results), and (c) the differences between (a) and (b).

6 Conclusions and Outlooks

This study suggests and investigates a new statistical multivariate seasonal forecasting approach for total water storage (TWS), which uses sea surface temperature and rainfall data alone to estimate TWS changes over West Africa. The proposed ICA/ARX approach does not directly simulate the complex physical process of ocean-land-atmosphere, but instead, it statistically learns the relationships between main known physical processes of the region (such as teleconnections and the soil-precipitation feedback) and uses it to predict TWS (the parameter of interest). The successful implementation of the proposed ICA/ARX approach relies on the proper selection of TWS indicators and avoiding over-parametrization of the model, data homogeneity and a learning phase that contains a thorough range of processes and impacts. Therefore, the dependence of the statistical model on the climate characteristics of the calibration period is often referred as a lack of model robustness. To investigate this issue, we performed a numerical validation, which showed that the seasonal forecast of TWS is close to TWS that is actually measured by GRACE, see Table 4. Both forward and backward predictions indicate that the proposed approach provides relatively stable large-scale seasonal TWS forecast over West Africa. We also carried out an extensive uncertainty analysis and were able to show that the predictability skills of the model is stable. However, due to the estimated uncertainties, the results might not be significant after about one year of forecast. We would like to mention here that the proposed method is only able to provide predictions of total water storage, therefore, hydrological modeling would still be required to partition TWS into different compartments. Since the prediction method relies on the relationships between SST and rainfall as indicators of TWS, we expect that the ICA/ARX method, with its current parameterization, is most appropriate to be used over those regions that exhibit strong interactions between ocean-atmosphere and land water storage changes, which is the case over West Africa.

Since the proposed method is trained on GRACE products, it provides relatively coarse resolution TWS maps. The approach also assumes that the spatial pattern of TWS changes remains stationary within the two years of the forecast. Before application, one should therefore analyze whether this assumption is fulfilled for different time frames. This can be achieved by applying the ICA technique to TWS time series of different length, and evaluating whether the dominant spatial patterns appear indeed invariant. Another issue is that the training of the ARX model was performed based on six years of the data. Since SST and rainfall are available for a longer period (e.g., for TRMM, since 1998), one could use TWS outputs of models for the time before 2002.8 and extend the training period. Addressing the impact of such extension in terms of the quality of the ARX coefficients and the consistency of the model-derived TWS with GRACE-TWS requires further research.

Our numerical results lead us to the hope that the presented statistical method could be helpful for filling the current gaps of the GRACE products (once every 162 days, with one or two months of data is missing) and a possible gap period between GRACE and its successor mission GRACE-FO at least for certain regions such as West Africa. Another application of the presented approach could be the generation of near-real time GRACE forecasts. Product latency time of GRACE fields is currently two to three months, while using the suggested approach, one

could be able to forecast GRACE total water storage maps immediately as soon as rainfall and sea surface temperature data become available. Such near-real time predictions could be used for various drought/flood monitoring applications. Generating total water storage predictions, close to GRACE products, would also be possible by calibrating and/or assimilating GRACE products in hydrological models. Models improved in this way could then be used to simulate total water storage. Examples of such implementations can be found in studies, e.g., Zaitchik et al. (2008), Werth et al. (2009b), Houborg et al. (2012), and Xie et al. (2012). The computational load of such approaches is however much more heavier than with the proposed statistical ICA/ARX approach.

Acknowledgement

The authors would like to thank M.J. Rycroft (EiC) and anonymous reviewers for their useful comments, which considerably improved this paper. We also thank S. Nahmani (LAboratoire de Recherche en Géodésie, France) for his detailed comments on the earlier version of this study. We are grateful for the GRACE, WGHM, TRMM, and SST data, as well as climate indices used in this study. E. Forootan and J. Kusche are grateful for the supports by the German Research Foundation (DFG), under the project DFG BAYES-G. The Ohio State University component of the research is supported by the NASA's Advanced Concepts in Space Geodesy Program (Grant No. NNX12AK28G), and by the Chinese Academy of Sciences/SAFEA International Partnership Program for Creative Research Teams (Grant No. KZZD-EW-TZ-05). The authors are grateful for the data used in this study. This is a TIGeR Publication no. xxx

References

Ahmed, M., Sultan, M., Wahr, J., Yan, E., Milewski, A., Sauck, W., Becker, R., & Welton, B. (2011). Integration of GRACE (Gravity Recovery and Climate Experiment) data with traditional data sets for a better understanding of the timedependent water partitioning in African watersheds. Journal of Geology, 41 (1), doi:10.1130/G31812.1.

Ali, A., & Lebel, T. (2009). The Sahelian standardized rainfall index revisited, Int. J. Climatol., 1714(December 2008), 1705-1714, doi:10.1002/joc.

Boone, A., Decharme, B., Guichard, F., de Rosnay, P., Balsamo, G., Beljaars, A., Chopin, F., Orgeval, T., Polcher, J., Delire, C., Ducharne, A., Gascoin, S., Grippa, M., Jarlan, L., Kergoat, L., Mougin, E., Gusev, Y., Nasonova, O., Harris, P., Taylor, C., Norgaard, A., Sandholt, I., Ottlé, C., Poccard-Leclercq, I., Saux-Picart, S., & Xue, Y. (2009). The AMMA Land Surface Model Intercomparison Project (ALMIP), Bull. Am. Meteorol. Soc., 90(12), 1865-1880, doi:10.1175/2009BAMS2786.1.

Chen, J.L., Wilson, C. R., Tapley, B.D., Longuevergne, L., Yang, Z.L., & Scanlon, B.R. (2010). Recent La Plata basin drought conditions observed by satellite gravimetry, J. Geophys. Res., 115(D22), 1-12, doi: 10.1029 /2010JD014689.

Crétaux, J.-F., Jelinski, W., Calmant, S., Kouraev, A., Vuglinski, V., Bergé Nguyen, M., Gennero, M.-C., Nino, F., Abarca Del Rio, F., Cazenave, A., & Maisongrande, P. (2011). SOLS: a lake database to monitor in near real time water level and storage variations from remote sensing data. J. Adv. Space Res., 14971507, http://dx.doi.org/10.1016/j.asr.2011.01.004.

Diatta, S., & Fink, A.H. (2014). Statistical relationship between remote climate indices and West African monsoon variability. Int. J. Climatol. (2014), in-press, doi:10.1002/joc.3912.

Döll, P., Kaspar, F., & Lehner, B. (2003). A global hydrological model for deriving water availability indicators: model tuning and validation. Journal of Hydrology 270 (1-2), 105-134, doi:10.1016/S00221694(02)002834.

Douville, H., Conil, S., Tyteca, S. and Voldoire, a. (2006). Soil moisture memory and West African monsoon predictability: artifact or reality?, Clim. Dyn., 28(7-8), 723-742, doi:10.1007/s00382-006-0207-8.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. Ann. Statist. 7, 1-26.

Eicker, A., Schumacher, M., Kusche, J., Döll, P., & Müller Schmied, H. (2014). Calibration/data assimilation approach for integrating GRACE data into the WaterGAP Global Hydrology Model (WGHM) using an ensemble Kalman filter. Surveys in Geophysics, submitted.

Flechtner, F. (2007). GFZ Level-2 processing standards document for level-2 product release 0004. GRACE 327-743, Rev. 1.0.

Fleming, K., & Awange, J.L. (2013). Comparing the version 7 TRMM 3B43 monthly precipitation product with the TRMM 3B43 version 6/6A and BoM datasets for Australia. Australian Meteorological and Oceanographic Journal (in press).

Forootan, E., Didova, O., Schumacher, M., Kusche, J., & Elsaka, B. (2014). Comparisons of atmospheric mass variations derived from ECMWF reanalysis and operational fields, over 2003 to 2011. Journal of Geodesy, 88 (5), 503-514, doi: 10.1007/s00190-014-0696-x.

Forootan, E., Didova, O., Kusche, J., & Löcher, A. (2013). Comparisons of atmospheric data and reduction methods for the analysis of satellite gravimetry observations, J. Geophys. Res. Solid Earth, 118, doi:10.1002/ jgrb.50160.

Forootan, E., Awange, J., Kusche, J., Heck, B., & Eicker, A. (2012). Independent patterns of water mass anomalies over Australia from satellite data and models. Journal of Remote Sensing of Environment, 124, 427-443, doi:0.1016/j.rse.2012.05.023.

Forootan, E., & Kusche, J. (2013). Separation of deterministic signals, using independent component analysis (ICA). Stud. Geophys. Geod. 57, 17-26, doi:10.1007/s11200-012-0718-1.

Forootan, E., & Kusche, J. (2012). Separation of global time-variable gravity signals into maximally independent components. Journal of Geodesy, 86 (7), 477-497, doi:10.1007/s00190-011-0532-5.

Giannini, A., Saravanan, R., & Chang, P. (2003). Oceanic forcing of Sahel rainfall on interannual to interdecadal time scales, Science, 302, 10271030.

Giannini, A., Biasutti, M., Held, I.M., & Sobel, A.H. (2008). A global perspective on African climate. Climatic Change 90, 359-383, doi:10.1007/ s10584-008-9396-y.

Grippa, M., Kergoat, L., Frappart, F., Araud, Q., Boone, A., de Rosnay, P., Lemoine, J.-M., Gascoin, S., Balsamo, G., Ottle, C., Decharme, B., Saux-Picart, S., & Ramillien, G. (2011). Land water storage variability over West Africa estimated by Gravity Recovery and Climate Experiment (GRACE) and land surface models. Water Resources Research 47, W05549, doi:10.1029/2009WR008856.

Güntner, A., Stuck, J., Döll, P., Schulze, K., Merz, B. (2007). A global analysis of temporal and spatial variations in continental water storage. Water Resources Research 43, (W05416), doi:10.1029/2006WR005247.

Hansen, J.W., Mason, S.J., Sun, L. & Tall, A. (2011). Review of seasonal climate forecasting for agriculture in sub-Saharan Africa, Expl Agric. 47 (2), pp. 20524047 (02), 205-240, : http://dx.doi.org/10.1017/ S0014479710000876.

Heim, R.R. (2002). A review of twentieth-century drought indices used in the United States. Bull. Amer. Meteor. Soc., 83, 1149-1165.

Houborg, R., Rodell, M., Li, B., Reichle, R. & Zaitchik, B.F. (2012). Drought indicators based on model-assimilated Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage observations, Water Resour. Res., 48, doi:10.1029/2011WR011291.

Huffman, G., & Bolvin, D. (2012). TRMM and other data precipitation data set documentation. Mesoscale Atmospheric Processes Laboratory, NASA Goddard Space Flight Center and Science Systems and Applications, Inc.

Ilin, A., Valpola, H., & Oja, E. (2005). Semiblind source separation of climate data detects El Niño as the component with the highest interannual variability. In: Proceedings of the International Joint Conference on Neural Networks (IJCNN 2005), Montréal, Québec, Canada (2005) 1722-1727.

Kaplan, A., Kushni, Y., Cane, M.A., & Blumenthal, M.B. (1997). Reduced space optimal analysis for historical data sets: 136 years of Atlantic sea surface temperatures. Journal of Geographical Research, 102 (C13), 27835-27860.

Koster, R.D., Dirmeyer, P.A., Guo, Z., Bonan, G., Chan, E., Cox, P., Gordon, C.T., Kanae, S., Kowalczyk, E., Lawrence, D., Liu, P., Lu, C-H., Malyshev, S., McAvaney, B., Mitchell, K., Mocko, D., Oki, T., Oleson, K., Pitman, A., Sud, Y.C., Taylor, C.M., Verseghy, D., Vasic, R., Xue, Y., & Yamada, T. (2004). Regions of strong coupling between soil moisture and precipitation. Science 20 August 2004: 305 (5687), 1138-1140, doi:10.1126/science.1100217.

Kusche, J. (2007). Approximate decorrelation and non-isotropic smoothing of time-variable GRACE-type gravity field models. Journal of Geodesy, 81 (11), 733-749, doi:10. 1007/s00190-007-0143-3.

Kusche, J., Schmidt, R., Petrovic, S., & Rietbroek, R. (2009). Decorrelated GRACE time-variable gravity solutions by GFZ, and their validation using a hydrological model. Journal of Geodesy, 83, 903-913, doi:10.1007/s00190-009-0308-3.

Lebel, T., Cappelarer, B., Galle, S., Hanan, N., Kergoat, L., Levis, S., Vieux, B., Descroix, L., Gosset M., Mougin, E., Peugeot, C., & Séguis, L. (2009). The AMMA-CATCH studies in the Sahelian region of West-Africa: an overview. Journal of Hydrology, 375, 313.

Long, D., Scanlon, B.R., Longuevergne, L., Sun, A.Y., Fernando, D.N. & Save, H. (2013). GRACE satellite monitoring of large depletion in water storage in response to the 2011 drought in Texas, Geophys. Res. Lett., 40(13), 3395-3401, doi:10.1002/grl.50655.

Ljung, L. (1987). System Identification - Theory for the User. Prentice Hall, Englewood Cliffs, N.J. 2005GL023316.

Mayer-Gürr, T., Eicker, A., Kurtenbach, E. (2010). ITG-GRACE 2010 unconstrained monthly solutions. http://www.igg.uni-bonn.de/apmg/

Mohino, E., Rodríguez-Fonseca, B., Mechoso, C.R., Gervois, S., Ruti, P., & Chauvin, F. (2011). Impacts of the tropical Pacific/Indian Oceans on the seasonal cycle of the West African monsoon. J. Clim., 24, 3878-3891, doi:10.1175 /2011JCLI3988.1.

Nahmani, S., Bock, O., Bouin, M-N., Santamaría-Gómez, A., Boy, J-P., Collilieux, X., Métivier, L., Panet, I., Genthon, P., de Linage, C., & Wöppelmann, G. (2012). Hydrological deformation induced by the West African Monsoon: Comparison of GPS, GRACE and loading models. Journal of Geophysical Research, 117, B05409, doi:10.1029/2011 JB009102.

Nicholson, S.E. (2000). The nature of rainfall variability over Africa on time scales of decades to millenia. Global and Planetary Change 26, 137-158.

Nicholson, S.E., and Coauthors (2003). Validation of TRMM and other rainfall estimates with a high-density gauge dataset for West Africa. Part I: Validation of GPCC Rainfall Product and Pre-TRMM Satellite and Blended Products. Journal of Applied Meteorology, 42, 1337-1354.

Omondi, P., Awange, J.L., Ogallo, L.A., Okoola, R.A., & Forootan, E. (2012). Decadal rainfall variability modes in observed rainfall records over East Africa and their relations to historical sea surface temperature changes. Journal of Hydrology, 464-465, 140156, doi: http://dx.doi.org/ 10.1016/ j.jhydrol.2012.07.003.

Preisendorfer, R. (1988). Principal component analysis in Meteorology and Oceanography. Elsevier: Amsterdam, 426 pages, ISBN:0444430148.

Reynolds, R.W., Rayne, N.A., Smith, T.M., Stokes, D.C., & Wang, W. (2002). An improved in situ and satellite SST analysis for climate. J. Clim. 15, 16091625.

Reager, J.T., Famiglietti, J.S. (2013). Characteristic mega-basin water storage behavior using GRACE, Water Resour. Res., 49, 33143329, doi:10.1002/ wrcr.20264.

Redelsperger, J-L., Thorncroft, Ch.D., Diedhiou, A., Lebel, T., Parker, D.J., & Polcher, J. (2006). African Monsoon multidisciplinary analysis: an international research project and field campaign. Bull. Amer. Meteor. Soc., 87, 1739-1746. doi:http ://dx.doi.org/10.1175/ BAMS-87-12-1739.

Rietbroek, R., Brunnabend, S.E., Dahle, C., Kusche, J., Flechtner, F., Schröter, J., & Timmermann, R. (2009). Changes in total ocean mass derived from GRACE, GPS, and ocean modeling with weekly resolution. Journal of Geophysical Research, 114, C11004, doi:10.1029/ 2009JC005449.

Rietbroek, R., Fritsche, M., Dahle, C., Brunnabend, S-E., Behnisch, M., Kusche, J., Flechtner, F., Schröter, J., & Dietrich, R. (2014). Can GPS-derived surface loading bridge a GRACE mission gap? Surv Geophys, in-press, doi:10.1007/s10712-013-9276-5.

Rodell, M., Houser, P.R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, K., Arsenault, C.-J., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J.K., Walker, J.P., Lohmann, D., & Toll, D. (2004). The Global Land Data Assimilation System. Bulletin of the American Meteorological Society, 85 (3), 381-394.

Rodríguez-Fonseca, B., Janicot, S., Mohino, E., Losada, T., Bader, J., Caminade, C., Chauvin, F., Fontaine, B., García-Serrano, J., Gervois, S., Joly, M., Polo, I., Ruti, P., Roucou, P. & Voldoire, A. (2011). Interannual and decadal SST-forced responses of the West African monsoon, Atmos. Sci. Lett., 12(1), 67-74, doi:10.1002/asl.308.

Saji, N.H., Goswami, B.N., Vinayachandran, P. N., & Yamagata, T. (1999). A dipole mode in the tropical Indian Ocean. Nature, 401, 360363, doi:10.1038/43854.

Scanlon, B.R., Faunt, C.C., Longuevergne, L., Reedy, R.C., Alley, W.M., McGuire, V.L. & McMahon, P.B. (2012). Groundwater depletion and sustainability of irrigation in the US High Plains and Central Valley., Proc. Natl. Acad. Sci. U. S. A., 109(24), 9320-5, doi:10.1073/pnas.1200311109.

Schmidt, R., Flechtner, F., Meyer, U., Neumayer, K.-H., Dahle, Ch., König, R., & Kusche, J. (2008a). Hydrological signals observed by the GRACE satellites. Surveys in Geophysics, 29 (4-5), 319-334.

Schmidt, R., Petrovic, S., Güntner, A., Barthelmes, F., Wünsch, J., Kusche, J. (2008b). Periodic components of water storage changes from GRACE and global hydrology models. Journal of Geophysical Research: Solid Earth, 113:B08419.

Schuol, J., Abbaspour, K. C., Yang, H., Srinivasan, R. & Zehnder, A.J.B. (2008). Modeling blue and green water availability in Africa, Water Resour. Res., 44, doi:10.1029/2007 WR006609.

Schuol, J., & Abbaspour, K.C. (2006). Calibration and uncertainty issues of a hydrological model (SWAT) applied to West Africa. Adv. Geosci., 9, 137-143.

Speth, P., Christoph, M., & Diekkrüger, B. (2011). Impacts of global change on the hydrological cycle in West and Northwest Africa. Speth, Peter; Christoph, Michael; Diekkrüger, Bernd (Eds.). Springer Berlin Heidelberg, 675 pages. ISBN:3642129560

Tapley B, Bettadpur S, Watkins M, & Reigber C. (2004). The gravity recovery and climate experiment: Mission overview and early results. Geophys Res Lett 31. http://dx.doi .org/10.1029/2004GL019920

van Dijk, A.I.J.M., Peña-Arancibia, J.L., Wood, Eric F., Sheffield, J., & Beck, H.E. (2013). Global analysis of seasonal streamflow predictability using an ensemble prediction system and observations from 6192 small catchments worldwide. Water Resources Research, 49 (5), 27292746, DOI: 10.1002/wrcr.20251.

von Storch, H., & Navarra, A. (1999). Analysis of climate variability. Springer, 342 p, ISBN 978-3-540-66315-7.

Wahr, J., Molenaar, M., & Bryan, F. (1998). Time variability of the Earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE. Journal of Geophysical Research, 103 (B12), 30205-30229, doi:10.1029/98JB02844.

718 Wang, K., & Dickinson, R.E. (2012). A review of global terrestrial evapotranspiration:
719 observation, modeling, climatology, and climatic variability. Reviews of Geophysics, 50
720 (2), doi:10.1029/2011RG000373.
721 Werth, S., Güntner, A., Schmidt, R., & Kusche, J. (2009a). Evaluation of GRACE filter
722 tools from a hydrological perspective. Geophysical Journal International, 179, 1499-1515.
723 http://dx.doi.org/10.1111/j.1365-246X.2009.04355.x.
724 Werth, S., Güntner, A., Petrovic, S., & Schmidt, R. (2009b). Integration of GRACE mass
725 variations into a global hydrological model. Earth and Planetary Science Letters 277(1),
726 166-173.
727 Westra, S., Brown, C., Lall, U., & Sharma, A. (2007). Modeling multivariable hydrological
728 series: Principal component analysis or independent component analysis? Water Resources
729 Research, 43 (6), W06429, doi:10.1029/2006WR005617.
730 Westra, S., Sharma, A., Brown, C., & Lall, U. (2008). Multivariate streamflow forecast-
731 ing using independent component analysis. Water Resources Research, 44 (2), W02437,
732 doi:10.1029/2007WR006104.
733 Xie, H., Longuevergne, L., Ringler, C., & Scanlon, B.R. (2012). Calibration and evaluation
734 of a semi-distributed watershed model of Sub-Saharan Africa using GRACE data, Hydrol.
735 Earth Syst. Sci., 16(9), 3083-3099, doi:10.5194/hess-16-3083-2012.
736 Zaitchik, B.F., Rodell, M., & Reichle, R.H. (2008). Assimilation of GRACE terrestrial
737 water storage data into a land surface model: results for the Mississippi River Basin, J.
738 Hydrometeor., 9 (3), 535-548, doi:10.1175/2007JHM951.1.

Appendix A (Computational Details of Total Water Storage Fields)

In order to prepare the data sets for analysis, the following processing steps were applied.

- The GRACE Level-2 data that are used here, are derived in terms of fully normalized spherical harmonic (SH) coefficients of the geopotential fields (Flechtner, 2007). Firstly, the fields were augmented by the degree-1 term from Rietbroek et al. (2009) in order to include the variation of the Earth's center of mass with respect to a crust-fixed reference system.
- GRACE SHs at higher degrees are affected by correlated noise and are, therefore, smoothed by applying the DDK2 decorrelation filter (Kusche et al., 2009). Werth et al. (2009a) found that the DDK2-filtered GRACE solutions are generally in good agreement with the output of global hydrological models. However, GRACE solutions are also contaminated by errors due to incomplete reduction of short-term mass variations by de-aliasing models (Forootan et al., 2013, 2014). We found that the impact of atmospheric de-aliasing errors on the GRACE-derived TWS over West Africa is negligible (see atmospheric errors over the Niger Basin in Forootan et al., 2014).
- GRACE DDK2 filtered solutions up to degree and order 120 were then used to generate the global TWS values according to the approach of Wahr et al. (1998).
- Similar to the GRACE products above, the DDK2 filter was applied to the gridded WGHM-TWS data set in order to preserve exactly the same spectral content as with the filtered GRACE products.
- After filtering, all data sets were converted to $0.5^\circ \times 0.5^\circ$ grids similar to the WGHM-TWS outputs.
- From each data set, a rectangular region that includes West Africa (latitude between 0° to $25^\circ N$ and longitude between -20° to $10^\circ E$) was selected.

Lake Volta (see Fig. A1) is one of the largest man-made reservoirs in the world, created by the Akosombo Dam, which holds back the water for generating hydroelectric power (for details see Speth et al., 2011). Satellite altimetry observations indicate a sharp increase of water level since mid 2007, where much of the excess water resulted from heavy rainfall within the catchment (Crétaux et al., 2011). This introduces an artificial TWS anomaly located over the lake, which is removed to avoid its misinterpretation as a part of subsurface TWS changes. The equivalent water height (EWH) change of Volta was computed by assuming a grid mask representing a unit change in EWH of 1 mm over the entire lake surface and zero elsewhere. The grid mask has been converted into a set of spherical harmonic coefficients up to degree 120 and subsequently filtered using the same DDK2 filter used for filtering the original GRACE-TWS data. Then, each field was scaled using the lake height time-series (in mm) derived from the results of Crétaux et al. (2011). The averaged storage changes derived from GRACE-TWS (from GFZ) and altimetry are shown in Fig. A1. Both altimetry and GRACE GFZ-TWS indicate an increase of water storage within the lake. The amplitude of the signal derived from GRACE GFZ-TWS is larger than that of the altimetry likely since GRACE-TWS also reflects the groundwater signal of the surrounding area of the lake. For the lake area, we estimate a TWS increase of $2.95 \pm 1.32\ km^3.yr^{-1}$, during 2003

to 2010. The time series of Lake Volta water storage changes were then removed from GRACE-TWS fields (including both GFZ and ITG2010).

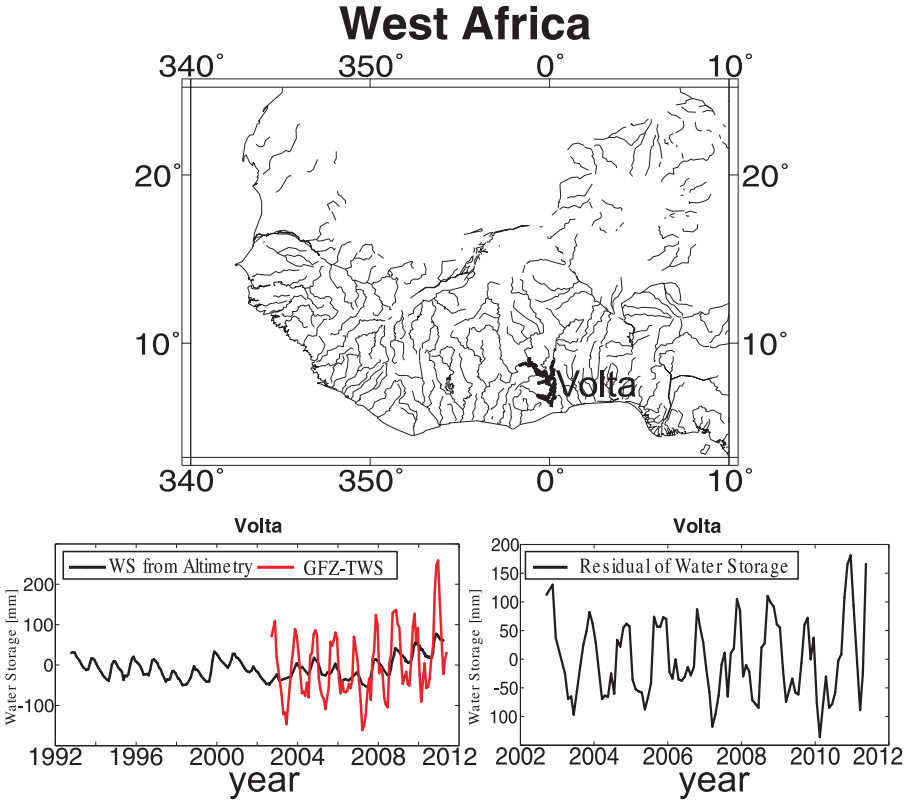

Figure A1: Overview of water storage changes of Volta Lake. The graph on top shows the location of the lake, while that of the bottom-left compares the averaged contribution of Volta lake level changes (derived from altimetry in black) with the averaged TWS variations derived from GRACE (GFZ-TWS products, in red). The bottom-right graph shows the GRACE GFZ-TWS signal after removing the water storage signal of Lake Volta.

In order to compare the signal strength over the region, the signal root-mean-square value (RMS) and the linear trend of the three mentioned TWS data sets (GFZ, ITG2010 and WGHM) are computed for the period January 2003 to August 2009, in which the three data sets were available (see Fig. A2). From the RMS, one concludes that all the three data sets show a strong variability over the tropical and the Gulf of Guinea coastal regions. The computed linear trends, however, are not consistent. Particularly, GRACE-derived TWS changes show a mass gain over Volta Lake, which we remove from the GRACE-TWS fields before performing decomposition. Removing such artificial anomaly is necessary, since otherwise the amplitude of TWS forecast over the lake will be overestimated.

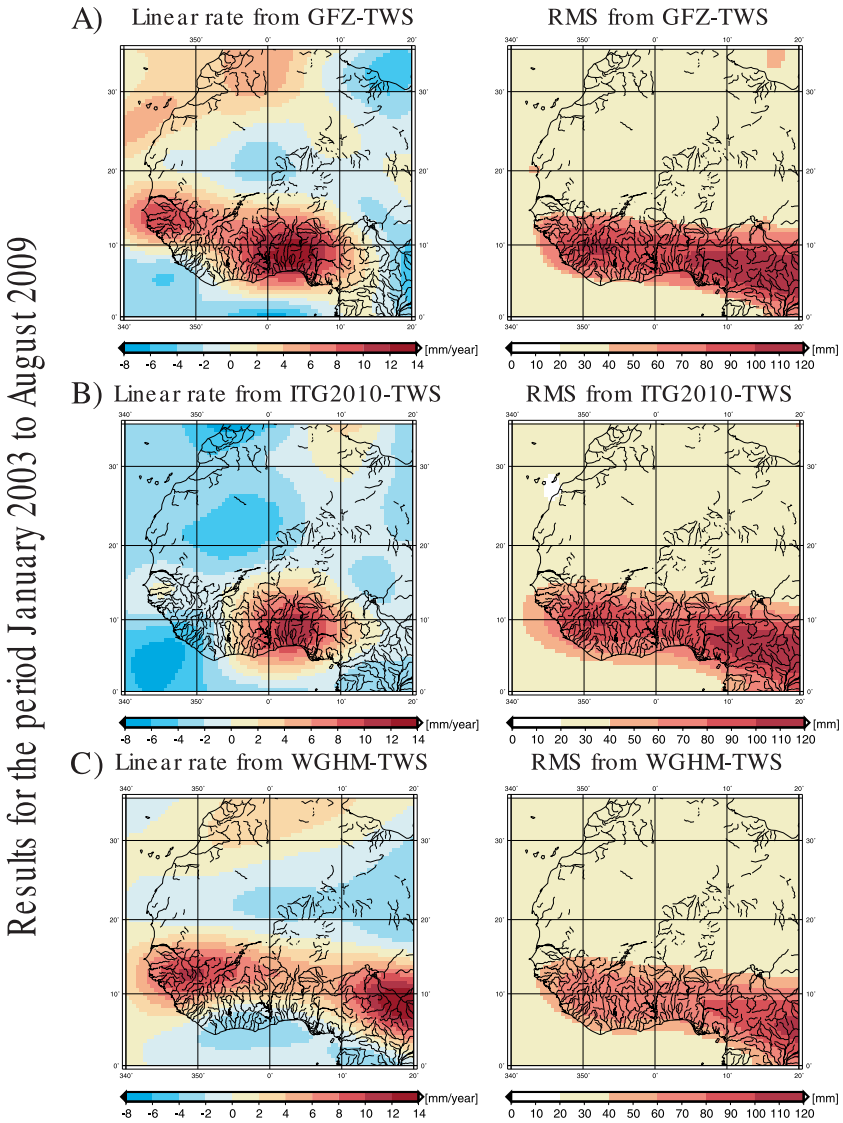

Figure A2: Comparing the signal variability (RMS) and linear trends of TWS data used in this study after smoothing using the Kusche et al. [2009]'s DDK2 filter. (A) TWS data of GRACE GFZ RL04, (B) TWS data of GRACE ITG2010, and (C) TWS of WGHM.

Appendix B (Details of ICA and ARX Methods)

This appendix provides details of computations regarding to the methodology described in Section 3.

The ICA Computations

ICA decomposition is performed here by applying a 2-step algorithm (Forootan and Kusche, 2012) on the available data sets, where step 1 consists of data decorrelation using principal component analysis (PCA). In step 2, the j-dominant components of PCA are rotated to be as independent from each other as possible. Storing the available data in a $n \times p$ data matrix $\mathbf{X}$, after removing their temporal mean, where n is the number of months and p is the number of grid points, ICA decomposes $\mathbf{X}$ as

$$\mathbf{X} \simeq \mathbf{X}_j = \bar{\mathbf{P}}_j \mathbf{R}_j \mathbf{\Lambda}_j \mathbf{R}_j^T \bar{\mathbf{E}}_j^T . \tag{B1}$$

In Eq. (B1), $\bar{\mathbf{P}}_j \mathbf{\Lambda}_j \bar{\mathbf{E}}_j^T$ is derived from the PCA decomposition of $\mathbf{X}$ in step 1. Therefore, $\mathbf{\Lambda}_j$ is an $j \times j$ diagonal matrix that stores the singular values arranged with respect to the magnitude, $\bar{\mathbf{E}}_j$ $(j \times p)$ contains the corresponding unit-length spatial eigenvectors, $\bar{\mathbf{P}}_j$ $(n \times j)$ contains the associated normalized temporal components, and $j < n$ is the number of retained dominant modes (Preisendorfer, 1988). The orthogonal rotation matrix $\hat{\mathbf{R}}_j$ $(j \times j)$ is defined in step 2, so that it rotates PCs and make them as statistically independent as possible. The method equals to temporal ICA (Forootan and Kusche, 2012), which is simply called ICA in the paper. Considering Eqs. (1) and (2), $\mathbf{Y}$ and $\mathbf{U}$ are equivalent to $\bar{\mathbf{P}}\mathbf{R}$, while $\mathbf{A}$ and $\mathbf{B}$ are equivalent to $\mathbf{\Lambda}\bar{\mathbf{E}}\mathbf{R}$. An optimum $\mathbf{R}$ was found by digonalization of the fourth-order cross cumulants of the dominant temporal components $\bar{\mathbf{P}}$ (see details in Forootan and Kusche, 2012).

For properly selecting the subspace dimension j or j', we used a Monte Carlo approach which simulates data from a random distribution $\mathbf{N}(\mathbf{0}, \mathbf{\Sigma})$, with $\mathbf{\Sigma}$ containing the column variance of $\mathbf{X}$. The null hypothesis is that $\mathbf{X}$ is drawn from such a distribution (see also Preisendorfer, 1988, pages 199 to 205). To apply the rule, 100 time series realizations of $\mathbf{N}(\mathbf{0}, \mathbf{\Sigma})$ are generated, their eigenvalues computed and placed in decreasing order. The 95th and 5th percentile of the cumulative distribution are then plotted (red lines in Fig. B1). Eigenvalues from the actual data sets that are above the derived confidence boundaries are unlikely to result from a data set consisting of only noise. To estimate the uncertainties of the eigenvalues, we randomly selected a subsample of $\mathbf{X}$ and applied PCA, then selected another subsample and repeated this operation 200 times. This approach follows the 'bootstrapping' method as presented, e.g., in Efron (1979) and yields uncertainty estimates (see error-bars in Fig. B1). The repeat number of 200 is chosen experimentally to be sure that the distribution of the estimated eigenvalues is independent from the selections of the subsamples.

To illustrate what we describe above, Fig. B1 shows the eigenvalue spectrum of the centered time series of GRACE GFZ-derived TWS, SST and rainfall computed using PCA. The significance levels are shown by red lines and the error-bars show the uncertainties of eigenvalues. The eigenvalues above the red lines are statistically significant. The significant eigenvalues along with their orthogonal components are rotated towards independence using Eq. (B1) and interpreted in Section 4.

Based on the uncertainties of the PCA results (Fig. B1), in order to estimate the uncertainty of the ICs (Eq. (B1)), we generated 100 realizations of $\mathbf{X}$, reconstructed by $\bar{\mathbf{P}}_j$, $\mathbf{\Lambda}_j$, and $\bar{\mathbf{E}}_j$ along with 100 realizations of their errors. Then, applying Eq.

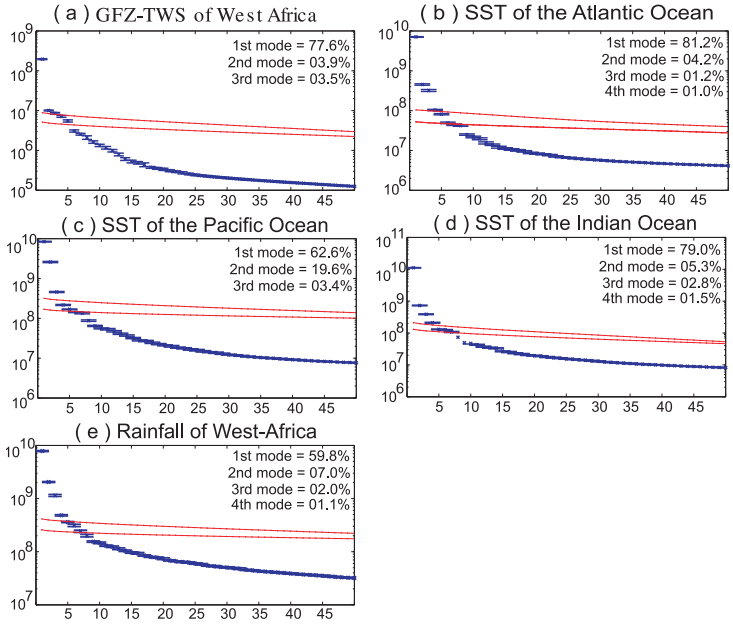

Figure B1: Eigenvalue results derived from implementing the PCA method on (a) time series of GRACE GFZ-TWS maps, (b) maps of SST over the Atlantic, (c) maps of SST over the Pacific, (d) maps of SST over the Indian Ocean basins, and (e) West-African rainfall maps from TRMM. Uncertainties are shown by error-bars around each eigenvalues. The red lines correspond to the significant test, described in Appendix B. The variance fractions of the dominant eigenvalues are represented in each graph.

(B1) to the realizations allows the estimation of uncertainties (see, e.g., error-bars in Figs. 1, 2, and 3).

The projection of the data $\mathbf{X}$ onto the i'th spatial pattern of the ICA $\hat{\mathbf{p}}_i = \mathbf{X}\hat{\mathbf{e}}_i$, provides its corresponding temporal evolution

$$\hat{\mathbf{p}}_i(t) = \sum_{s=1}^{p} x(t,s)\ \hat{\mathbf{e}}_i(s), \tag{B2}$$

where t is time $(1, \cdots, n)$ and s is the number of grid points $(1, \cdots, p)$.

The ARX Computations

Considering Eq. (3) as the ARX model, the ARX-forecast requires two steps: (i) The coefficients $(a_1, \cdots, a_{n_a})$ and $(b_{q,1}, b_{q,2}, \cdots, b_{q,n_b}), q = 1, \cdots, m$ are estimated, e.g., using a least squares approach. This step is usually referred to as 'simulation' or 'training step' in literature (see, e.g., Ljung, 1987). Step (i) is performed under the assumption that the output and inputs up to the time $t = t_n - 1$ are known. Furthermore, the outputs and exogenous values on the right hand side

of Eq. (3) are not stochastic. To avoid negative indices, one might consider the observations $\mathbf{y}(t) = [y(t), y(t-1), \cdots, y(c)]^T$, where $c = max(n_a, n_b) + max(k_q) + 1$. Eq. (3) is expanded as

$$\mathbf{y} = \begin{bmatrix} -y(t-1) & \cdots & -y(t-n_a) & u_q(t-k_q) & \cdots & u_q(t-k_q-n_b+1) \\ -y(t-2) & \cdots & -y(t-n_a-1) & u_q(t-k_q-1) & \cdots & u_q(t-k_q-n_b) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -y(c-1) & \cdots & -y(c-n_a) & u_q(c-k_q) & \cdots & u_q(c-k_q-n_b+1) \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_{n_a} \\ b_{q,1} \\ \vdots \\ b_{q,n_b} \end{bmatrix} + \boldsymbol{\Xi}(t), \tag{B3}$$

where $q = 1, \cdots, m$ and $\boldsymbol{\Xi}(t) = [\xi(t), \xi(t-1), \cdots, \xi(c)]^T$. Eq. (B3) can be rewritten compactly as

$$\mathbf{y}(t) = \boldsymbol{\Phi}(t)\boldsymbol{\Theta} + \boldsymbol{\Xi}(t). \tag{B4}$$

The least squares estimation of the unknown coefficients is derived from

$$\hat{\boldsymbol{\Theta}} = \left(\boldsymbol{\Phi}(t)^T \boldsymbol{\Phi}(t)\right)^{-1} \boldsymbol{\Phi}(t)^T \, \mathbf{y}(t). \tag{B5}$$

The quality of the fit (η) can be assessed by computing the signal-to-noise ratio as

$$\eta = 1 - \frac{\mathbf{y}(t)^T \boldsymbol{\Phi}(t) \hat{\boldsymbol{\Theta}}}{\mathbf{y}(t)^T \mathbf{y}(t)}. \tag{B6}$$

The residual of the ARX model ($\hat{\boldsymbol{\Xi}}(t) = [\hat{\xi}(t), \hat{\xi}(t-1), \cdots, \hat{\xi}(c)]^T$) can be estimated as

$$\hat{\boldsymbol{\Xi}}(t) = \mathbf{y}(t) - \boldsymbol{\Phi}(t) \, \hat{\boldsymbol{\Theta}}. \tag{B7}$$

In step (ii), based on $\hat{\boldsymbol{\Theta}} = \begin{bmatrix} \hat{a}_1 & \ldots & \hat{a}_{n_a} & \hat{b}_{q,1} & \hat{b}_{q,2} & \cdots & \hat{b}_{q,n_b} \end{bmatrix}^T$, when the inputs $u_q(t)$ are known, one can forecast the output $\hat{y}(t_n)$ at time t_n using

$$\hat{y}(t_n) = -\sum_{i=1}^{n_a} \hat{a}_i y(t_n - i) + \sum_{q=1}^{m} \sum_{l=1}^{n_b} \hat{b}_{q,l} u_q(t_n - k_q - (l-1)). \tag{B8}$$

To estimate the uncertainty of the ARX simulation, using Monte Carlo sampling, we numerically generate several realizations of the ICs (described before). By inserting them into Eq. (3) and fitting ARX models, we are able to perform an error assessment of the fitted model up to the time t_n. For error estimation of the forecast (the ARX value at the time $t_n + 1$ and later), however, one should compute an accumulated error, since there is no observed value for the output y at time $t_n + 1$ and later.